

**IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MASSACHUSETTS**

NUANCE COMMUNICATIONS, INC.,

Plaintiff and Counterclaim
Defendant,

v.

OMILIA NATURAL LANGUAGE
SOLUTIONS, LTD.,

Defendant and Counterclaim
Plaintiff.

Case No. 1:19-CV-11438-PBS

**OMILIA NATURAL LANGUAGE SOLUTIONS, LTD.'S MEMORANDUM IN
SUPPORT OF ITS MOTION FOR PARTIAL SUMMARY JUDGEMENT FOR
THE INVALIDITY OF U.S. PATENT 6,999,925**

TABLE OF CONTENTS

I.	INTRODUCTION.....	1
II.	FACTUAL BACKGROUND.....	2
III.	LEGAL STANDARD	4
IV.	ARGUMENT.....	7
A.	The '061 Patent Is an Obviousness-Type Double Patenting Reference	7
B.	Claim 6 of the '061 Patent Renders Claim 1 of the '925 Patent Invalid for Obvious-Type Double Patenting.....	9
1.	Claim 1 of the '925 Patent Covers the Speech Recognition Adaptation Technique Claimed in the Earlier Claim 6 of the '061 Patent	9
2.	The Semantical Differences of Claim 1 of the '925 Patent Do Not Render it Patentably Distinct from the Earlier Claim 6 of the '061 Patent	11
C.	Claim 15 of the '061 Patent Renders Claim 14 of the '925 Patent Invalid for Obvious-Type Double Patenting.....	15
D.	Claim 6 of the '061 Patent Renders Claim 27 of the '925 Patent Invalid for Obvious-Type Double Patenting.....	16
1.	Applying Claim 6 of the '061 Patent to a Multilingual First Speech Recognizer Results in a Multilingual Second Speech Recognizer	18
2.	Adapting a First Speech Recognizer to a Dialect Is Conceptually the Same as Adapting a First Speech Recognizer to a Language	19
3.	Adaptation of a First Speech Recognizer to Create a New Multilingual Speech Recognizer Was Known in the Art	20
V.	CONCLUSION	20

TABLE OF AUTHORITIES

Page(s)**CASES**

<i>Amgen Inc. v. Hoffmann–LaRoche, Ltd.</i> , 580 F.3d 1340 (Fed. Cir. 2009).....	5
<i>Athletic Alternatives, Inc. v. Prince Mfg., Inc.</i> , 73 F.3d 1573 (Fed. Cir. 1996).....	4
<i>Boehringer Ingelheim Int’l GmbH v. Barr Laboratories, Inc.</i> , 592 F. 3d 1340 (Fed. Cir. 2010).....	6, 7
<i>Eli Lilly & Co. v. Barr Labs., Inc.</i> , 251 F.3d 955 (Fed. Cir. 2001).....	4, 6, 17
<i>Eli Lilly & Co. v. Teva Parenteral Meds., Inc.</i> , 689 F.3d 1368 (Fed. Cir. 2012).....	4
<i>Geneva Pharmaceuticals, Inc. v. GlaxoSmithKline PLC</i> , 349 F.3d 1373 (Fed. Cir. 2003).....	6, 7, 14, 15
<i>Geneva Pharms., Inc. v. Glaxosmithkline PLC</i> , 213 F. Supp. 2d 597 (E.D. VA 2002)	10
<i>Georgia-Pacific Corp. v. United States Gypsum Co.</i> , 195 F.3d 1322 (Fed. Cir. 1999).....	19
<i>Gilead Sciences Inc. v. Natco Pharma. Ltd.</i> , 753 F.3d 1208 (Fed. Cir. 2014).....	5, 7
<i>Immunex Corp. v. Sandoz, Inc.</i> , 395 F. Supp. 3d 366 (D. N.J. 2019)	5
<i>In re Basell Poliolefine Italia S.P.A.</i> , 547 F.3d 1371 (Fed. Cir. 2008).....	6
<i>In re Berg</i> , 140 F.3d 1428 (Fed. Cir. 1998).....	6
<i>In re Hubbell</i> , 709 F.3d 1140 (Fed. Cir. 2013).....	6, 7
<i>In re Lonardo</i> , 119 F.3d 960 (Fed. Cir. 1997).....	16

In re Longi,
759 F.2d 887 (Fed. Cir. 1984).....6

In re Vogel,
422 F.2d 438 (C.C.P.A 1970)13

Miller v. Eagle Mfg. Co.,
151 U.S. 186 (1894).....4, 7

Pfizer, Inc. v. Teva Pharms. USA, Inc.,
518 F.3d 1353 (Fed. Cir. 2008).....5

STATUTES AND RULES

35 U.S.C. § 2537

Fed. R. Civ. P. 56(c)4

MISCELLANEOUS

U.S. Patent 6,789,061 passim

U.S. Patent 6,999,925 passim

I. INTRODUCTION

Omilia Natural Language Solutions, Ltd. (“Omilia”) moves for summary judgment that the asserted claims 1, 14, and 27 of U.S. Patent 6,999,925 (the “’925 patent”) are invalid under the doctrine of obviousness-type double patenting based on U.S. Patent 6,789,061 (the “’061 patent”). The prohibition on double patenting exists to prevent situations just like this, where an inventor impermissibly extends a patent’s exclusivity by securing a second patent with a later expiration date on the same invention. A Statement of Undisputed Material Facts to which there is no issue to be tried and the declaration of Dr. Jordan Cohen support this motion.

The ’925 and ’061 patents were filed by, issued to, and initially owned by the International Business Machines Corporation (“IBM”) and have two common inventors. The ’061 patent expired in 2008 for failure to pay maintenance fees. Nuance Communications, Inc. (“Nuance”) bought the ’925 patent with hundreds of other IBM patents in December 2008. Nuance cannot overcome obviousness-type double patenting by filing a terminal disclaimer since the ’061 patent has expired and the patents are not commonly owned.

Obviousness-type double patenting is a question of law with underlying factual inquiries and focuses on whether the claims of the second patent are an obvious variant of the claims of the first patent. Here, the method for adapting a first speech recognizer to a second speech recognizer claimed in the earlier ’061 patent is covered by the later ’925 patent claims. The ’061 patent claims this method in terms of adapting the mathematical components of the recognizer, while the ’925 patent claims the same process in terms of adapting the functional structures of the recognizer. The differences in the ’925 patent are mere semantics or obvious variations of the ’061 patent. Summary judgment that the later issued ’925 patent is invalid for double patenting in view of the claims of the earlier ’061 patent is warranted.

II. FACTUAL BACKGROUND

The '061 patent was filed by IBM on behalf of the inventors Volker Fischer, Siegfried Kunzmann, and Claire Waast-Ricard on August 14, 2000. Statement of Undisputed Material Facts (“SUF”) at ¶¶1-2. The '061 patent claims a method and apparatus for automatically generating or adapting a second speech recognizer from a first speech recognizer. SUF at ¶5. The inventors assigned the '061 patent to IBM in September 2000, prior to its issuance on September 7, 2004. SUF at ¶¶3-4. The '061 patent would have expired on or around January 31, 2021, but expired on September 7, 2008 due to IBM’s failure to pay the required maintenance fee. SUF at ¶6.

On November 13, 2001, IBM filed a second patent application, which later issued as the '925 patent. SUF at ¶7. The '925 patent also claims a method and system for generating or adapting a second speech recognizer from a first speech recognizer. SUF at ¶13. Two inventors of the '061 patent—Volker Fischer and Siegfried Kunzmann—are also named as inventors of the '925 patent along with Eric-W. Janke and Jon Tyrrell. SUF at ¶¶2, 8-9. The '925 patent was assigned to IBM in October 2001 and issued on February 14, 2006. SUF at ¶¶10-11. The '925 patent will expire on or around October 12, 2023 and is not subject to a terminal disclaimer. SUF at ¶¶12, 16. Two years after the '925 patent issued, Nuance bought the patent along with hundreds of other IBM patents, executing the assignment on December 31, 2008. SUF at ¶15.

The applications for the patents were not before the same Examiners in the Patent Office and were prosecuted by different law firms. *See* '925 patent at 1, (74); '061 patent at 1, (74). The European patent application from which the '061 patent claims priority was cited in the specification of the '925 patent, but the '061 patent claims were not evaluated as an obviousness type double patenting reference during the prosecution of the '925 patent. *See* SUF at ¶ 14.

Both the '925 and '061 patents refer to IBM’s ViaVoice German language speech

recognizer and indicate the ViaVoice recognizer can be automatically modified or adapted using the methods the patents disclose. '925 patent at 9:26-10:4; '061 patent at 7:52-8:60. Both patents claim a method of adapting a first speech recognizer to a second domain-specific speech recognizer. '925 patent, cl. 1; '061 patent, cl. 1. The speech recognizer of the '061 patent is described as a set of states and a set of probability density functions. '061 patent, cl. 1 at 9:38-39. The claimed method of the '061 patent generates a second speech recognizer from a first recognizer by selecting a subset of the set of states and set of probability density functions in the first recognizer that are distinctive to a specific domain. '061 patent at 9:42-53. The '925 patent describes the same speech recognizer of the '061 patent as comprising an acoustic model made up of a decision network and corresponding phonetic context. '925 patent at 10:44-46, 10:48-51. The '925 patent claims a method of adapting the second recognizer by re-estimating the phonetic contexts of the first speech recognizer using domain-specific training data. '925 patent at 10:51-53. The '925 patent method covers the adaptation described in the '061 patent.

The specification of the '925 patent refers to the foreign priority European patent application of the '061 patent, EP99116684.4, and incorporates it by reference. '925 patent at 6:57-65, 8:58-9:3. It describes the methods of EP99116684.4 as a method for adapting a speech recognizer: by "squeezing," which involves selecting a subset of HMM states of the general purpose recognizer; and "pruning," which involves selecting a subset of probability density functions of the general purpose recognizer. '925 patent 8:58-9:3. Both such selections are made based on the domain to which the second speech recognizer is adapted. '925 patent at 9:2-3. The '925 patent incorporates the methods of EP99116684.4 by reference and goes on to describe three additional aspects of '925 patent method. '925 patent 9:2-23. Claim 1 of the '925 patent encompasses the method claimed disclosed in EP99116684.4 and claimed in the '061 patent.

On August 6, 2020, the Court construed the following '925 patent claim terms:

Term (claims)	Court's Construction
"a second language" (claims 12, 25, 27)	"a language other than the first language of the first domain/speech recognizer." (ECF No. 157 at 13)
"multi-lingual speech recognizer" (claims 12, 25, 27)	"a speech recognizer that is able to cover more than one language" (<i>Id.</i> at 14)
"generating a second acoustic model" (claim 27)	Consistent with the remaining independent claims as to require "re-estimating said first decision network and said corresponding first phonetic contexts." (<i>Id.</i> at 14)
"automatically generate/ing" (claims 1, 2, 14, 15)	Plain and ordinary meaning (<i>Id.</i> at 21)

On October 16, 2020, Nuance narrowed its asserted claims in the '925 patent to claims 1, 14, and 27.

III. LEGAL STANDARD

Summary judgment should be granted where there is no genuine issue as to any material fact and the moving party is entitled to judgment as a matter of law. Fed. R. Civ. P. 56(c); *Athletic Alternatives, Inc. v. Prince Mfg., Inc.*, 73 F.3d 1573, 1578 (Fed. Cir. 1996). Invalidity due to double patenting is a proper ground for summary judgment. *Eli Lilly & Co. v. Barr Labs., Inc.*, 251 F.3d 955, 968 (Fed. Cir. 2001). Obviousness-type double patenting is a question of law "premised on underlying factual inquiries." *Eli Lilly & Co. v. Teva Parenteral Meds., Inc.*, 689 F.3d 1368, 1376 (Fed. Cir. 2012).

"No patent can issue for an invention actually covered by a former patent, especially to the same patentee, although the terms of the claims may differ; that the second patent, although containing a broader claim, more generic in its character than the specific claims contained in the prior patent, is also void." *Miller v. Eagle Mfg. Co.*, 151 U.S. 186, 198 (1894). The rule

against double patenting exists to prevent an inventor from extending a patent's term of exclusivity by patenting later variations of the first-patented invention that are not patentably distinct. *Immunex Corp. v. Sandoz, Inc.*, 395 F. Supp. 3d 366, 408-09 (D. N.J. 2019) (citing *Applied Materials, Inc. v. Advanced Semiconductor Materials Am., Inc.*, 98 F.3d 1563, 1568 (Fed. Cir. 1996)). The judicially-created doctrine of obviousness-type double patenting furthers the statutory limitations by prohibiting a patent owner from obtaining an improper time-wise extension of its exclusive rights when "claims in separate applications or patents [] do not recite the 'same' invention, but nonetheless claim inventions so alike." *Amgen Inc. v. Hoffmann-LaRoche, Ltd.*, 580 F.3d 1340, 1352 (Fed. Cir. 2009) (citing *Perricone v. Medicis Pharm. Corp.*, 432 F.3d 1368, 1373 (Fed. Cir. 2005)). "[I]t is a bedrock principle of our patent system that when a patent expires, the public is free to use not only the same invention claimed in the expired patent but also obvious or patentably indistinct modifications of the invention." *Gilead Sciences Inc. v. Natco Pharma. Ltd.*, 753 F.3d 1208, 1214 (Fed. Cir. 2014). The doctrine of double patenting protects the public's right to use the claimed invention once the statutory monopoly period concludes by preventing a time-wise extension of the earlier patent's monopoly. *Id.*

The first step in an obviousness-type double patenting analysis is to determine whether the earlier expiring patent is an obviousness-type double patenting reference. The patents must have a common inventor or common ownership, *Immunex Corp.*, 395 F. Supp. 3d at 409, and the later issued patent seeks to improperly extend the monopoly on the invention of the earlier issued patent through a later expiration date, *Gilead Sciences Inc.*, 753 F.3d at 1214. The expiration dates, not the issuance dates, determine whether a double patenting defense exists. *Id.* at 1216-17.

Courts apply a two-part test to determine whether obviousness-type double patenting applies. *Pfizer, Inc. v. Teva Pharms. USA, Inc.*, 518 F.3d 1353, 1363 (Fed. Cir. 2008). The court

begins by comparing the claims of the earlier patent and the claims in the later patent and determines any differences. *See Eli Lilly*, 251 F.3d at 967. The court may also examine the specifications of the patents to ascertain any overlap in claim scope. *Geneva Pharmaceuticals, Inc. v. GlaxoSmithKline PLC*, 349 F.3d 1373, 1385 (Fed. Cir. 2003). Once any differences have been identified, the court determines whether those differences render the claims patentably distinct by applying a one-way test, asking whether the later patent's claims are anticipated or obvious over the earlier patent's claims. *Ely Lilly*, 251 F.3d at 968; *see also In re Basell Poliolefine Italia S.P.A.*, 547 F.3d 1371, 1375-1376 (Fed. Cir. 2008) (applying the "one-way test," except in "unusual circumstances"). A prior claim is anticipatory if it discloses every limitation of the later claimed invention either explicitly or inherently. *Ely Lilly*, 251 F.3d at 970. When the claims of the later patent encompass the claims of the earlier patent, the earlier patent anticipates the later issued patent and the relationship between the patents is one of genus and species. *See In re Berg*, 140 F.3d 1428, 1431 (Fed. Cir. 1998) (holding a later genus patent invalid as anticipated by obviousness-type double patenting over an earlier species patent).

To overcome an obviousness-type double patenting violation and permit a second patent on an obvious variation of a prior patent, the patentee of commonly owned patents may file a terminal disclaimer, disclaiming the term of the second patent that extends beyond that of the first patent. *In re Longi*, 759 F.2d 887, 894 (Fed. Cir. 1984); *see also In re Hubbell*, 709 F.3d 1140, 1148 (Fed. Cir. 2013). That remedy is not available where the first and second patent with common inventors are not commonly owned. *See In re Hubbell*, 709 F.3d at 1148-1149. That remedy is also not available after the expiration of the earlier patent over which claims have been found obvious. *Boehringer Ingelheim Int'l GmbH v. Barr Laboratories, Inc.*, 592 F.3d 1340, 1348 (Fed. Cir. 2010).

IV. ARGUMENT

The question in obviousness-type double patenting is what the patents claim. The specification gives life and meaning to the claims. *Geneva Pharmaceuticals, Inc.*, 349 F.3d at 1385. The '925 patent claims as drafted are broad enough to include the method claimed in the '061 patent. *See* Cohen Decl. at ¶¶35-43. The '061 patent claim 6 (and 15) is a species of the genus of the asserted '925 patent claims. Both patents claim a method for generating a second speech recognizer from a first. By generating a second speech recognizer through a process of re-estimation that selects a subset of states of the first recognizer, the '061 patent claims a method of re-estimation encompassed by the '925 patent claims. Cohen Decl. at ¶¶36-38, 58-62. In this instance, IBM when prosecuting the later-filed and issued '925 patent drafted claims encompassing those of its earlier-expiring '061 patent. The '925 patent is, therefore, void. *Miller*, 151 U.S. at 198.

A. The '061 Patent Is an Obviousness-Type Double Patenting Reference

The '925 patent and the '061 patent have two common inventors—Volker Fischer and Siegfried Kunzmann (SUF at ¶ 9)—and were both owned by IBM when the '925 patent issued on February 14, 2006 (SUF at ¶¶ 3, 10). The '925 patent expires on or around October 12, 2023, more than a year after the '061 patent's original January 31, 2021 expiration date and more than 12 years after its actual expiration in September 2008 for failure to pay maintenance fees. SUF at ¶¶ 6, 16. Thus, the '061 patent constitutes an obviousness-type double patenting reference against the '925 patent. *Gilead Sciences Inc.*, 753 F.3d at 1214. Because the patents are no longer commonly owned, Nuance may not file a terminal disclaimer of the '925 patent term to cure the deficiency. *See* 35 U.S.C. § 253; *see also In re Hubbell*, 709 F.3d at 1148-1149. Even if the patents were commonly owned, a terminal disclaimer could not cure the invalidity of the '925 patent because the '061 patent has expired. *Boehringer Ingelheim Int'l GmbH*, 592 F. 3d at 1348.

The '061 patent was not cited by the applicant during prosecution of the '925 patent and was not considered by the Examiner. '925 patent at 1, (56). IBM, the co-owner of both patents at the issuance of the '925 patent, did not file a terminal disclaimer disclaiming the term of the '925 patent extending beyond the term of the earlier expiring '061 patent. SUF at ¶ 12. Instead, IBM abandoned the '061 patent by failing to pay the required maintenance fees before selling the '925 patent to Nuance. SUF at ¶ 6.

Nevertheless, the specification of the '925 patent refers to the foreign priority European patent application of the '061 patent, EP99116684.4, and incorporates it by reference. '925 patent at 6:57-65, 8:58-9:3. It describes the methods of EP99116684.4 as “squeezing” and “pruning”, and creating recognizers that are distinctive of the specific domain. '925 patent 8:58-9:3. The inventors of the '925 patent described the methods of EP99116684.4 as an aspect of their broader method and '925 patent claim 1 encompasses the method of EP99116684.4, claimed in claim 6 of the '061 patent.

The '061 patent describes optionally collecting and using a small amount of training data for a particular application or domain to adapt the acoustic model. '061 patent at FIG. 1 (2) & (10), 4:67-5:5, 6:3-5, 7:3-20, 8:36-40. By selecting the subset of states, the corresponding phonetic contexts are also sub-selected, meaning that both the states and the phonetic contexts of the second speech recognizer are different from the first speech recognizer. Cohen Decl. at ¶¶61-62.

The '925 patent defines the re-estimation of phonetic contexts as a recalculation of the decision network and corresponding phonetic contexts (the speech recognizer’s “subword inventory”) by “insertion, deletion or adaptation of phones in their respective context.” '925 patent at 7:2-17. Most significantly, in distinguishing the previous V. Fischer methods (EP99116684.4), the '925 patent states:

This is considerably different from *just* “selecting” a subset of the general speech recognizer decision network and phonetic contexts or simply “enhancing” the decision network by making a leaf node an interior node by attaching a new sub-tree with new leaf nodes and further phonetic contexts.

’925 patent at 7:12-18 (emphasis added).

The key word is “*just*.” While the ’925 claimed method may include the insertion and modification of phones, it also includes the methods claimed in the ’061 patent of selecting certain phones and deleting others. The claims of the ’925 patent were drafted broadly and encompass a method that re-estimates the first speech recognizer decision network by only deleting phones, just as claimed in the ’061 patent. *See* Cohen Decl. at ¶¶35-43.

The supporting Cohen Declaration explains the language used in the claims of the ’061 and ’925 patents and discusses how each of the limitations of the asserted claims 1, 14, and 17 of the ’925 patent are found in claims 6 and 15 of the ’061 patent.

B. Claim 6 of the ’061 Patent Renders Claim 1 of the ’925 Patent Invalid for Obvious-Type Double Patenting

1. Claim 1 of the ’925 Patent Covers the Speech Recognition Adaptation Technique Claimed in the Earlier Claim 6 of the ’061 Patent

The earlier ’061 patent claim 6 contains every element recited in the later claim 1 of the ’925 patent. Both claims are directed to the same subject matter—adapting a first speech recognizer into a second speech recognizer. The ’925 patent describes this process through the abstract functional structure of the recognizer whereas the ’061 patent describes the same process through the mathematical elements that comprise the recognizer. *See* Cohen Decl. at ¶35. Both patents are based on IBM’s ViaVoice German speech recognizer. ViaVoice is the only speech recognizer used in the examples or referred to in both patents.

As explained in the declaration of Dr. Jordan Cohen, claim 6 of the ’061 patent, which is dependent on claims 1, 2, 4, and 5 and contains all their limitations, claims a species of what is

claimed in claim 1 of the '925 patent. The earlier claim 6 generates a second speech recognizer from a first speech recognizer by “selecting a subset of the set of states” and “probability density functions” of the first speech recognizer based on application-specific training data.

The later '925 patent claim 1 generates a second speech recognizer from a first speech recognizer in the same manner; however, the '925 patent describes this process as “re-estimating the decision network and phonetic contexts” of the first speech recognizer based on domain-specific training data. '925 patent 6:18-20. Re-estimation is defined in the '925 patent as the specific “insertion, deletion *or* adaptation of phones,” '925 patent 7:7-8 (emphasis added). Claim 1 of the '925 patent includes a method where phones are deleted only. As Dr. Cohen explains, the principle differences between the two claims are semantics. Claim 6 of the '061 patent recites the mathematical components that make up the functional elements described in claim 1 of the '925 patent. The earlier claim 6 of the '061 patent is to a species of what is claimed in the '925 patent, and therefore anticipates claim 1 of the '925 patent and renders its invalid under obviousness-type double patenting. *See* Cohen Decl. ¶¶36-43.

A person of ordinary skill in the art (POSA) would understand that the different terms used in the claims of the '061 and '925 patents describe the same concepts. Any differences between the claims here are merely differences in word choice and do not render the claims patentably distinct. *See Geneva Pharms., Inc. v. Glaxosmithkline PLC*, 213 F. Supp. 2d 597, 607 (E.D. VA 2002) (finding a later patent an obvious variant of an earlier patent in part because the “differences between the patents are works of semantics”).

2. The Semantical Differences of Claim 1 of the '925 Patent Do Not Render it Patentably Distinct from the Earlier Claim 6 of the '061 Patent

The Cohen Declaration compares each of the limitations of claim 1 of the '925 patent and claim 6 of the '061 patent and explains how the specific limitations correspond and, as a result, describe the same method. The semantical differences of the claims are summarized below.

The '925 patent claim 1 describes a speech recognizer with an acoustic model, which includes a decision network and phonetic context. '925 patent cl. 1 at 10:44-46. The decision network is organized as a decision tree. *Id.* at 10:54-56. The first speech recognizer is modified using domain-specific training data to create a second speech recognizer by creating a second, or updated, acoustic model. *Id.* at 10:48-53. The modification is accomplished by “re-estimating” the first acoustic model’s decision network and its corresponding phonetic contexts. *Id.* The language of the claim provides “the number of nodes in the second decision network is not fixed by the number of nodes in the first decision network.” *Id.* at 57-58. This re-estimation of the phonetic structure may include adding, deleting or modifying phones in the decision network, which changes the number of nodes. *Id.* at 7:2-12. By virtue of this re-estimation of the phonetic structure, the number of nodes in the second speech recognizer is not fixed by the number of nodes in the first speech recognizer. *Id.* at 10:56-59. The '925 patent describes “each terminal node of the adapted (i.e. generated) decision network defines a context dependent, single state Hidden Markov Model for the specialized speech recognizer.” *Id.* at 8:41-44. The re-estimation also includes partitioning the training data using the first speech recognizer, i.e. running the training data through the first speech recognizer. *Id.* at 10:59-61, 7:54-61.

The '061 patent claim 6 also claims a first speech recognizer adapted to a second speech recognizer using domain-specific training data. '061 patent cl. 1 at 9:36-40. It adapts the phonetic structure of the acoustic model of the first recognizer but describes it based on the mathematical

elements that make up the acoustic model. Cohen Decl. at ¶42. Specifically, the acoustic model is made up of a set of states (and corresponding set of probability density functions), and these states comprise a decision network and phonetic context. *Id.*; '061 patent at 4:1-27, 4:48-53. The states and their corresponding probability density functions correspond to the nodes in the decision network claimed in the '925 Patent. Cohen Decl. at ¶42. The states in the claims of the '061 patent are the terminal/leaf nodes of a decision tree in a speech recognizer. Cohen Decl. at ¶¶33, 42; '061 patent at 4:16-19. The '061 patent claims a process for re-estimating the phonetic contexts of the first acoustic model by selecting a subset of the first speech recognizer corresponding to the training data, i.e. deleting states not needed for the training data. '061 patent at 9:42-55. By deleting unused states and probability density functions, the number of nodes of the decision network is changed. Cohen Decl. at ¶42. As a result in claim 6 of the '061 patent, the number of nodes in the second decision network is not fixed by the number of nodes in the first decision network, just as in claim 1 of the '925. Cohen Decl. at ¶42. Finally, the '061 patent claims associating the training data with the corresponding states in the decision network and then selecting those states within the training data. '061 patent at 10:13-17; Cohen Decl. at ¶42.

The differences between the functional structures of the '925 patent, “decision network and phonetic context,” and the mathematical elements of those structures in the '061 patent, “set of states and probability density functions,” are semantics and of no patentable significance. Cohen Decl. at ¶¶36-43. A decision network as used in both the '925 and '061 patents comprises nodes organized to represent the relationship between the phones (i.e. phonetic context). Each terminal node of the decision network represents a phone in context. Cohen Decl. at ¶33. These terminal nodes correspond to a state as claimed in the '061 patent, described in the '925 patent at 8:41-44 and understood in the art. Cohen Decl. at ¶¶33, 42. By claiming a set of states, the '061 patent

therefore claims a decision network and its corresponding phonetic contexts. Cohen Decl. at ¶¶42, 47-53. This differing vocabulary refers to the same concepts in both patents.

While the first and second decision networks “utilize a phonetic decision tree to perform speech recognition operations” in claim 1 of the ’925 patent (’925 patent at 10:55-56), the ’061 patent claim 6 claims the same as a set of states of the speech recognizer. These states in the ’061 patent refer to HMM states or leaves. *See* Cohen Decl. at ¶¶33, 51. The ’061 patent repeatedly describes its set of states as part of conventional, general purpose speech recognizer. *See, e.g.*, ’061 patent at 4:58-5:5; 7:42-50, Fig. 1. That same structure is used in the specialized second speech recognizer that is an adapted version of the first recognizer. These speech recognizers make use of a binary decision network in speech recognition, which a POSA would have understood refers to a decision tree. Cohen Decl. at ¶¶66; ’061 patent at 4:16-19. A POSA would likewise understand that the first recognizer and second recognizer use “binary decision networks” or set of states, which “separates phonetic contexts” into different leaves or states. ’061 patent at 4:11-23. A set of states are not only HMMs states, but those states are leaf nodes in the decision tree used for speech recognition. Cohen Decl. at ¶¶33, 66. These states are, therefore, part of a “binary decision network” (’061 patent at 4:17-18), which describes a decision tree. Cohen Decl. at ¶¶66.

Re-estimation in the ’925 patent encompasses the selection of a subset of states as claimed in the ’061 patent claim 6. The ’925 patent defines re-estimation as adding, deleting or modifying phones in the respective context. ’925 patent at 7:5-8; Cohen Decl. at ¶¶41, 60. The ’061 patent re-estimates the second decision network by deleting nodes, i.e. selecting a subset of states. ’061 patent, cl. 6; Cohen Decl. at ¶¶42, 61. The genus in the ’925 patent is therefore anticipated by the species claimed in claim 6 of the ’061 patent. Cohen Decl. at ¶¶62; *see In re Vogel*, 422 F.2d 438,

442-43 (C.C.P.A 1970) (finding a later issued genus anticipated by the earlier issued species where the term “meat” literally read on the term “pork”).

While claim 6 does not include an express limitation that the nodes of the decision network are not fixed, ('925 patent at 10:56-59), claim 6 recites generating a set of states for the second speech recognizer by **selecting** a subset of the states of the first speech recognizer. Cohen Decl. at ¶71. Each state represents a node, or leaf, of a decision network. Because claim 6 requires the selection of a subset of the states from the first speech recognizer, the number of nodes, i.e. states, necessarily cannot be fixed by the number of nodes in the first speech recognizer as only some of the nodes, i.e. states, will be selected for the second speech recognizer. That the number of nodes of the decision network are not fixed is, therefore, a necessary result of practicing the method of claim 6 of the '061 patent. *See Geneva Pharmaceuticals, Inc.*, 349 F.3d at 1385.

Claim 6 of the '061 patent recites the same “domain-specific” limitations as in the '925 patent when it states that the second recognizer is “tailored to a specific application” and that the selection of a subset of states “exploits” or is based upon “application specific training data.” The terms “application” in claim 6 of the '061 patent and “domain” (“domain-specific”) in claim 1 of the '925 patent are interchangeable. Cohen Decl. at ¶¶55-56, 63. As described in the '061 patent, an “application” dependent space “is like, e.g., a certain dialect or domain.” '061 patent at 4:37-38. This is consistent with the common meaning of “domain” in the '925 patent. '925 patent at 6:5-8 (“A domain might refer to a certain language, a multitude of languages, a dialect or a set of dialects, a certain task area or set of task areas.”). Thus, the '925 patent’s use of “domain” is the same as the use of “application” in the '061 patent and claim 6 of the '061 patent describes the same process of adapting a second recognizer for a specific domain using domain-specific training data.

Claim 6 of the '061 patent recites the same “partitioning of training data” limitation as present in the '925 patent when it recites that “selecting the subset of states comprises associating a multitude of speech frames of the training data with the correct states of the first speech recognizer.” '061 patent, cl. 6. Claim 6 of the '061 patent requires that part of the training data (speech frames) are associated with the correct states of the first speech recognizer’s set of states. There is no meaningful difference between this “associating” of training data in claim 6 of the '061 and the “partitioning” of training data in claim 1 of the '925 patent. Cohen Decl. at ¶¶ 72-74. Claim 6 of the '061 patent, therefore, uses the term “associating” to describe the step of partitioning the training data claimed in claim 1 of the '925 patent.

When understood in light of the '061 patent specification, the differences in the claims are semantics. The method of claim 6 of the '061 patent is a species of the genus of claim 1 of the '925 patent rendering the broader genus of the '925 patent anticipated by claim 6 of the '061 patent. The practice of claim 6 of the '061 patent would infringe claim 1 of the '925 patent. Cohen Decl. at ¶¶ 34-43. *See Geneva Pharmaceuticals, Inc.*, 349 F.3d at 1384 (finding an earlier species anticipates a later genus rendering the later claims patentable indistinct).

C. Claim 15 of the '061 Patent Renders Claim 14 of the '925 Patent Invalid for Obvious-Type Double Patenting

Claim 14 of the '925 patent is almost identical to claim 1 of the '925 patent except for its additional recitation of a “machine-readable storage, having stored thereon a computer program” that causes the machine to perform the recited verbatim method of claim 1 of the '925 patent. *Compare* '925 patent at claim 1, *with* '925 patent at claim 14. Claim 15 of the '061 patent likewise identically claims an apparatus for automatically generating a second speech recognizer using the same method described in claim 6 of the '061 patent. Thus, for the same reasons that the limitations of claim 6 of the '061 patent render claim 1 of the '925 patent invalid, claim 15 of the

'061 patent renders claim 14 of the '925 patent invalid for obviousness-type double patenting.

As described for claim 6 of the '061 patent and claim 1 of the '925 patent, claim 15 of the '061 patent recites the same limitations as in the later claim 14 of the '925 patent. *See supra* Section IV.B. The references in claim 14 of the '925 patent to generic computer elements (computer storage and processing) do not render it patentably distinct from the apparatus of claim 15 of the '061 patent. *See In re Lonardo*, 119 F.3d 960, 968 (Fed. Cir. 1997) (rejecting argument that “the method of using the device would not have been obvious over a claim to the device,” as there was no “patentable distinction between the method of using the device and the device itself”). A POSA would understand that the apparatus for implementing the method of claim 15 of the '061 patent includes computer storage and a processor. Cohen Decl. at ¶¶75-76; *see also* '061 patent at 2:55-3:8 (describing the invention as applied on “any kind of computer system” and embedded on a “computer program product”).

D. Claim 6 of the '061 Patent Renders Claim 27 of the '925 Patent Invalid for Obvious-Type Double Patenting

Claim 27 of the '925 is an obvious variant of claim 6 of the '061 patent. *See* Cohen Decl. at ¶¶77-80. Claim 27 of the '925 patent recites a subset of the limitations discussed in connection with claim 1 of the '925 patent but adds two limitations: (1) receipt of domain specific training data; and (2) the first domain comprising a first language and the second domain comprising a second language resulting in the second speech recognizer being a multilingual speech recognizer.

In its claim construction ruling, the Court construed “generating” in claim 27 of the '925 patent to require “re-estimating said first decision network and said corresponding first phonetic contexts,” which is an express limitation of claim 1 of the '925 patent. Dkt. No. 157 at 22. In view of this claim construction, and for the same reasons discussed above in relation to claim 1 of the '925 patent, claim 6 of the '061 patent recites an adaptation process covered by claim 27 of the

'925 patent.

Neither of the two additional limitations of claim 27 of the '925 patent render it patentably distinct from claim 6 of the '061 patent. Using the adaptation method described in claim 6 to incorporate different languages or to further adapt a multilingual recognizer are obvious variants of the method described in claim 6 of the '061 patent. Multilingual recognizers were well known as were adaptations to languages of recognizers. Claim 27 of the '925 patent is an obvious variant of applying the adaptation method of claim 6 of the '061 patent to the multilingual recognizers of the prior art.

Claim 27 claims "receiving domain-specific training data of a second domain." '925 patent at 14:13-14. Claim 6 does not explicitly claim the receipt of domain-specific training data; however, it is clear from claim 6 that domain-specific training data must be received to practice the claimed method. The selection of states in claim 6 "exploits application-specific training data," ('061 patent at 10:9-12), and "associat[es] a multitude of speech frames of the training data with the correct states of the first speech recognizer," (*id.* at 10:13-17). To practice claim 6, one must receive the training data in order to exploit it and associate it with states of the first speech recognizer. Cohen Decl. at ¶¶81; *see also* '061 patent at 4:16-19, 6:9-16, 6:30-33, 7:3-7. The inclusion in claim 27 of the '925 patent of a limitation that requires "receiving domain specific training data" does not render the claim patentably distinct from claim 6 of the '061 patent and in any event, this limitation is inherent in claim 6. *Eli Lilly & Co.*, 251 F.3d at 970-71 (finding one limitation of the later patent inherently claimed by the earlier patent and obviousness-type double patenting renders the claims of the later patent invalid).

'925 patent claim 27 further recites that the adaptation of the decision network and phonetic context involves a first recognizer of a first domain, comprising at least a first language, and that

the resulting second recognizer is multilingual and covers the first domain and a second domain comprising at least a second language. This is little more than an obvious application of the method recited in claim 6 of the '061 patent to a known and conventional use for speech recognizers. Multi-lingual speech recognizers were well known in the art. Applying the method of claim 6 to a multilingual first recognizer would yield a multilingual second speech recognizer.

1. Applying Claim 6 of the '061 Patent to a Multilingual First Speech Recognizer Results in a Multilingual Second Speech Recognizer

A POSA would have understood that applying claim 6 of the '061 patent to a multilingual recognizer could have resulted in an adapted multilingual recognizer as claimed by claim 6 of the '925 patent. Cohen Decl. at ¶¶84-85. Multilingual speech recognizers were known in the art as exemplified by the references cited during prosecution of '925 patent, for example Schultz et al., “*Language Adaptive LVCSR through Polyphone Decision tree Specialization*,” Workshop on Multi-lingual Interoperability in Speech Technology (MIST 1999), Leusden, The Netherlands, Sep. 1999, pp. 85-90. Cohen decl. at ¶86 (providing examples of prior art describing multilingual recognizers). If the domain-specific training data contains data from both languages, the selection of a subset of states would still select states for both language. Cohen Decl. at ¶85. The adaptation by selecting a subset of states described in claim 6 of the '061 patent to known multilingual recognizers could result in a second recognizer that is capable of covering more than one language (i.e. “multilingual” as construed in this case). Cohen Decl. at ¶¶84-85; Dkt. No. 157 at 13. Nuance’s own expert agrees. In a declaration in support of Nuance’s claim construction relating to the '925 patent Dr. Livescu acknowledged that a POSA would have understood that the first recognizer in claim 27 could be multilingual as could the training materials used to adapt the first recognizer. Dkt. No. 86 at ¶36 (“A POSITA would understand that based on [the] language [of claim 27], both the first domain and the second domain can include multiple languages, but need

not do so.”); SUF at ¶18. Thus, applying the adaptation claimed in claim 6 of the ’061 patent to known multilingual recognizers was an obvious variant and renders claim 27 of the ’925 patent invalid. *See Georgia-Pacific Corp. v. United States Gypsum Co.*, 195 F.3d 1322, 1328 (Fed. Cir. 1999) (dependent claims invalid for obviousness-type double patenting where the variation between the claims was well-known in the prior art).

2. Adapting a First Speech Recognizer to a Dialect Is Conceptually the Same as Adapting a First Speech Recognizer to a Language

A POSA would have understood that the claim 6 method exemplified in the ’061 patent to adapt a German language speech recognizer to the Austrian dialect could be extended to adapt a recognizer to recognize distinct languages. At the time the ’925 patent application was filed, multilingual speech recognizers not only existed but inventors were developing ways to provide better recognition accuracy for cross- and multilingual system. One such example, as outlined in both the ’925 and ’061 patents, was IBM’s ViaVoice system. *See* ’925 patent at 3:29-33; ’061 patent at 1:40-45.

The ’061 patent explicitly describes that the methods it claims “can be applied to move the acoustic model parameters of a HMM based speech recognizer to an application dependent acoustic space, like, e.g., a certain dialect or domain,” 4:35-38. The ’061 patent contains an example of applying the claimed methods to IBM’s ViaVoice German language speech recognizer and re-estimating it with a small amount of speech data from Austrian speakers (i.e., domain-specific training data) to create a dialect-specific speech recognizer from a large vocabulary speech recognition system for general German. ’061 patent 7:57-8:67. Speech recognizers recognize sounds and the same concepts that apply to the recognition of different sounds in different dialects apply to the recognition of the different sounds in different languages. Thus, the ’061 patent’s method of adapting based on domain specific training data could be applied to a language just as

it could be applied a dialect. Cohen Decl. at ¶¶87-88. Claim 27 of the '925 patent is, therefore, an obvious variant of claim 6 of the '061 patent.

3. Adaptation of a First Speech Recognizer to Create a New Multilingual Speech Recognizer Was Known in the Art

Not only was the “multilingual” use obvious based on the '061 patent itself, but at the time of the '925 patent it was also known that one could adapt a first speech recognizer to add a second language by deleting, modifying and adding phonetic contexts. Cohen Decl. at ¶¶89-94. For example, prior art to the '925 patent, U.S. Pat. No. 6,912,499, to Sabourin et al., for a “Method and Apparatus for Training a Multilingual Speech Model Set” filed August 31, 1999 (“Sabourin”) described a method for creating a second multilingual speech recognizer by adapting a first speech recognizer. SUF at ¶17; Cohen Decl. at ¶¶89-90; Sabourin, cl. 1. Sabourin discloses automatically identifying the existing phones of a first recognizer and adding any new phones that may be needed for the new second language based on the training data from that second language. Cohen Decl. at ¶90; Sabourin at 6:45-7:8. Sabourin describes the resulting speech recognizer as multilingual. Cohen Decl. at ¶90; Sabourin at 2:20-22. The '061 patent not only contemplates changing phonetic contexts for dialects, but that the accuracy of recognizers could be further improved by “investing” in new phonetic contexts. Cohen Decl. at ¶¶88, 91. A POSA would be motivated to combine Sabourin with the method of '061 patent claim 6 with a reasonable expectation of success to adapt the phonetic contexts of speech recognizers to yield a multilingual recognizer as claimed in claim 27 of the '925 patent. Cohen Decl. at ¶¶91-92. Claim 27 of the '925 patent is an obvious variant of claim 6 of the '061 patent in light of the prior art. Cohen Decl. at ¶94.

V. CONCLUSION

For the forgoing reasons, the Court should grant Omilia’s motion for summary judgment that claims 1, 14, and 27 of the '925 patent are invalid over claims 6 and 15 of the '061 patent.

Dated: November 25, 2020

Respectfully Submitted,

By /s/ Daniel S. Sternberg

Kevin C. Adam (SBN 684955)
Daniel S. Sternberg (SBN 688842)
WHITE & CASE LLP
75 State Street, 24th Floor
Boston, MA 02109
(617) 979-9300
Kevin.adam@whitecase.com
Daniel.sternberg@whitecase.com

Of Counsel:
Dimitrios Drivas (*admitted pro hac vice*)
Raj Gandesha (*admitted pro hac vice*)
John Padro (*admitted pro hac vice*)
Stefan Mentzer (*admitted pro hac vice*)
WHITE & CASE LLP
1221 Avenue of the Americas
New York, NY 10020-1095
(212) 819-8286
ddrivas@whitecase.com
rgandesha@whitecase.com
john.padro@whitecase.com
smentzer@whitecase.com

Hallie Kiernan (*admitted pro hac vice*)
WHITE & CASE LLP
3000 El Camino Real
Two Palo Alto Square, Suite 900
Palo Alto, CA 94306
(650) 213-0300
hallie.kiernan@whitecase.com

*Counsel for Omilia Natural Language
Solutions, Ltd.*

CERTIFICATE OF SERVICE

I hereby certify that counsel of record who are deemed to have consented to electronic service are being served on November 25, 2020, with a copy of this document via the Court's CM/ECF system per Local Rule CV-5.4 (c). Any other counsel of record will be served by First Class U.S. Mail on this same date.

/s/ Daniel S. Sternberg
Daniel S. Sternberg